DOCUMENT RESUME

ED 440 111 TM 030 714

AUTHOR Lee, Yew Jin; Linacre, John M.; Yeoh, Oon Chye
TITLE KIDMAP--A Diagnostic Tool for Teachers.
PUB DATE 1999-00-00
NOTE 7p.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Diagnostic Tests; Elementary Secondary Education; Evaluation Methods; Foreign Countries; *Item Response Theory; *Student Evaluation; Test Items; *Test Results
IDENTIFIERS *Kidmaps

ABSTRACT

While assessment is the bread and butter of the teaching profession, its practitioners usually do not extend analysis of test responses beyond simple measures such as facility or discrimination indices in classical test theory. Item response theory (IRT) has much to offer but its nonintuitive content and difficulty make it a formidable obstacle in teachers' use of this assessment methodology. B. Wright, R. Mead, and L. Ludlow (1980) have devised a useful IRT diagnostic tool called KIDMAP that is increasingly being popularized in local school contexts. KIDMAP provides feedback on individual students' responses to test items. Strengths and weaknesses of the use of KIDMAP are briefly discussed. (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 111

KIDMAP - A Diagnostic Tool for Teachers

Lee, Y. J., Linacre, J. M., & Yeoh, O. C.

Mr Lee Yew Jin, &
Dr Yeoh Oon Chye,
School of Science,
National Institute of
Education,
469 Bukit Timah Road
Singapore 259756

Prof Linacre, John, M.
MESA Psychometric
Laboratory,
University of
Chicago5835 Kimbark
Ave, Chicago, Il 60637,
USA

BEST COPY AVAILABLE

TM030714

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Lee Yew Jin

ABSTRACT

While assessment is the bread-and-butter of the teaching profession, its practitioners usually do not extend analysis of test responses beyond simple measures such as facility or discrimination indices in classical test theory. Item response theory (IRT) has much to offer but its non-intuitive content and difficulty make it a formidable obstacle in teachers' use of this assessment methodology. Wright, Mean and Ludlow (1980) have devised a useful IRT diagnostic tool called a KIDMAP which is increasingly being popularised in local school contexts. Strengths and weaknesses of the use of KIDMAPS are briefly discussed in this paper.

Introduction

While assessment is the bread-and-butter of the teaching profession, its practitioners usually do not extend analysis of test responses beyond simple measures such as facility or discrimination indices in classical test theory. A possible reason is the lack of time among teachers who often have to cover a very tight curriculum and thus rely on simple test analysis done by a hand-held calculator. Item response theory (IRT) has much to offer but its non-intuitive content and difficulty make it a formidable obstacle in teachers' use of this assessment methodology.

Many students in Singapore primary and secondary schools participate in quizzes conducted by the University of New South Wales in Australia yearly. These highly popular quizzes solely using the multiple choice question (MCQ) format, are conducted in content areas such as English, Science, Mathematics and Computing. Feedback is provided for each question attempted by the individual. Certificates of Merit and Distinction are awarded to the top scorers from the school, and in the nation as a whole.

What was interesting to the first author who is a practising high school science teacher was that a useful IRT diagnostic tool called a KIDMAP (Wright, Mead & Ludlow, 1980) was being used to provide feedback on individual students' responses for items in the quiz. The KIDMAP provided was a simplified version which converted students' ability levels in logit units into true scores while items retained relative positions according to their difficulty on the same linearized scale. The KIDMAP was explained to students and their parents in clear and simple language.

It is not known to what extent KIDMAPS are used in everyday teaching in Singapore schools beyond those provided by the Australian quizzes. We consider this to be a great loss as KIDMAPS can provide valuable information on possible areas of weakness in knowledge or misconceptions in the individual as illustrated below in Figure 1. This KIDMAP was generated from the QUEST computer programme for one-parameter Rasch analysis in an achievement test for evolution amongst high school teachers in Singapore (Lee & Yeoh, 1998).

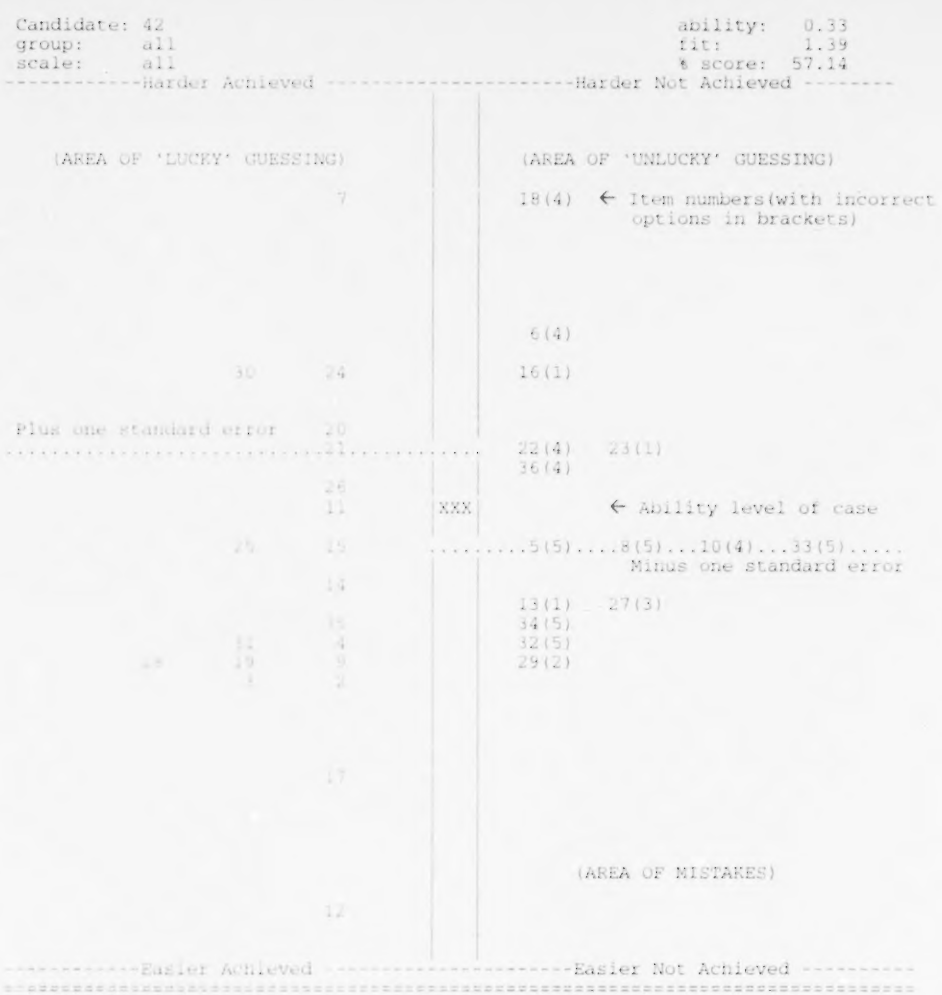

Figure 1. KIDMAP of Case number 42 which exhibited probable guessing behaviour in answering the achievement test.

Here, person/case number 42 had five items which he/she should have been able to answer correctly in the Easier Not Achieved (ENA) category but did not. It might have been due to 'unlucky' guessing or some unexplainable factor. If a consistent pattern in content area occurred in the ENA category of the KIDMAP on other tests, it is likely that this is an area of weakness or indicative of possible misconceptions.

Having a number of questions in both the Harder Not Achieved (HNA) and ENA categories should alert one to the possibility of random guessing behaviour by the student. As this is a linearized scale, items distant from the ability level of the student represent those which are considered very difficult (eg. items 7 and 18) or very easy (eg. items 12 and 17) in Figure 1.

In preparing students for the GCE O Level examinations, the first author had felt that feedback for students using KIDMAPs had rather limited utility in diagnosing misconceptions. One main reason was that there were too few items in each content subsection (usually a total of 40 MCQ in the GCE O Level science examinations) to be able to reach any definite conclusions about students' misconceptions. Most students had in fact

BEST COPY AVAILABLE

attributed their items in the ENA categories to carelessness rather than a genuine misconception or lack of knowledge. It is felt that this should not detract from using KIDMAPs in a teacher's assessment repertoire. Feedback for MCQ tests(if any at all) for local students has generally been confined to percentiles and mean scores derived from a certain population. However, a KIDMAP would provide instead an individualised record of a student's answering pattern and its psychological effects on esteem should not be discounted.

And while it is undisputed that the KIDMAP is a useful diagnostic tool for teachers, it is certainly not easy to produce one! Some knowledge of IRT computer programmes are usually necessary such as QUEST or TITAN (Doig, 1992). It is not known if easy to use, stand-alone software is available for teachers which bypasses the somewhat abstract commands of QUEST and other programmes. There is certainly a need to provide one for busy teachers who have little time to spare on learning computer languages. What is now highlighted is a procedure which can be utilised to produce a rough KIDMAP using nothing more than a hand-held calculator (Michael Linacre, personal communication, 24/07/98).

Procedure:

1. Find number of wrong and correct responses to each test item.
2. Item location = log(number of wrong answers/number of right answers) to that item.

This gives the item location with 0 at 50% success. The middle of the items is about 50% on the test, item quartiles correspond to roughly 25% and 75%.

3. Re-rank if necessary the items according to difficulty.
4. Plot these items on the y-axis of a graph, x=0.
5. Calculate the ability level of the individual which is
   = log(number of right answers/number of wrong answers) and indicate on this on the graph by drawing a horizontal line across that point.
6. Items **above** the ability level and are

a) incorrect : place a tick on the right of the item (the Harder Not Achieved category)
b) correct : place a tick on the left of the item (the Harder Achieved category)

7. Items **below** the ability level and are

a) incorrect : place a tick on the right of the item (the Easier Not Achieved category)
b) correct : place a tick on the left of the item (the Easier Achieved category)

8. The standard error for person is =

$$\text{Squareroot of } \left(\frac{\text{count of right + count of wrong answers}}{\text{(count of right answers * count of wrong answers)}} \right)$$

Now, all four quadrants of the KIDMAP should be visible and analysis proceeds as usual.

Example:

Number of MCQ items in test =30

a) For item number 1 (rather difficult item), there were 10 correct answers and 30 incorrect ones. Item location for item 1= log (30/10) = 0.48 units

b) For item number 2 (rather easy item), there were 30 correct answers and 10 incorrect ones. Item location for item 2= log (10/30) = -0.48 units

c) For item number 3 , there were 15 correct answers and 15 incorrect ones. Item location for item 3= log (15/15) = 0 units

d) Item locations are thus calculated for the rest of the 27 items as shown in the previous three steps.
e) Respondent A obtained 25 marks in the test, the ability level is thus=log (25/5) = 0.70 units
f) Standard error is thus

$$\frac{\sqrt{(25+5)}}{(25*5)} = 0.49 \text{ units}$$

The ability levels for the rest of the respondents can be calculated accordingly.

There are however some disadvantages in this rough and ready analysis. The KIDMAP output produced by a programme based on IRT like QUEST provides other important information, for example reliability values and MNSQ values which help pinpoint misfitting individuals quickly. It has been our experience that most people do conform to the norm by having items mainly in the Harder Not Achieved and Easier Achieved categories. In the first author's experience, a 40 item MCQ biology test often had not more than 10 pupils who showed KIDMAPs worthy of further discussion with the instructor from a sample of about 130 pupils. Indeed, with many items and cases, this manual method of producing a KIDMAP is certainly tedious compared to analysis done by a computer programme. Without pupil assistance to help produce the four quadrants after the teacher has calculated the item positions, this method is almost impossible to perform for a class size of 40 which is common in Singapore.

References

Doig, B. A. (1992). Self Scoring Kidmaps. ERIC Document Reproduction Number ED346146.

Lee, Y.J. and Yeoh, O.C. (1998). Biological evolution : a tough nut to crack for biology teachers in Singapore? Journal of Popular Measurement, 1,

Wright, B. D., Mead, R. J. and Ludlow, L. H. (1980). Kidmap. Research Memorandum No. 29. University of Chicago.